

# 2D unfolding linearity tests

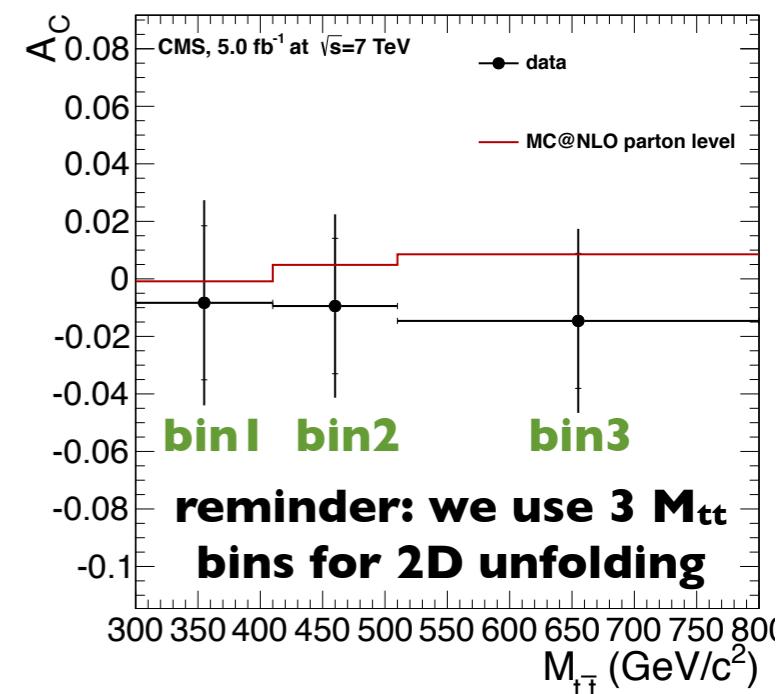
Jacob Linacre, for TOP-12-010 authors

# Introduction

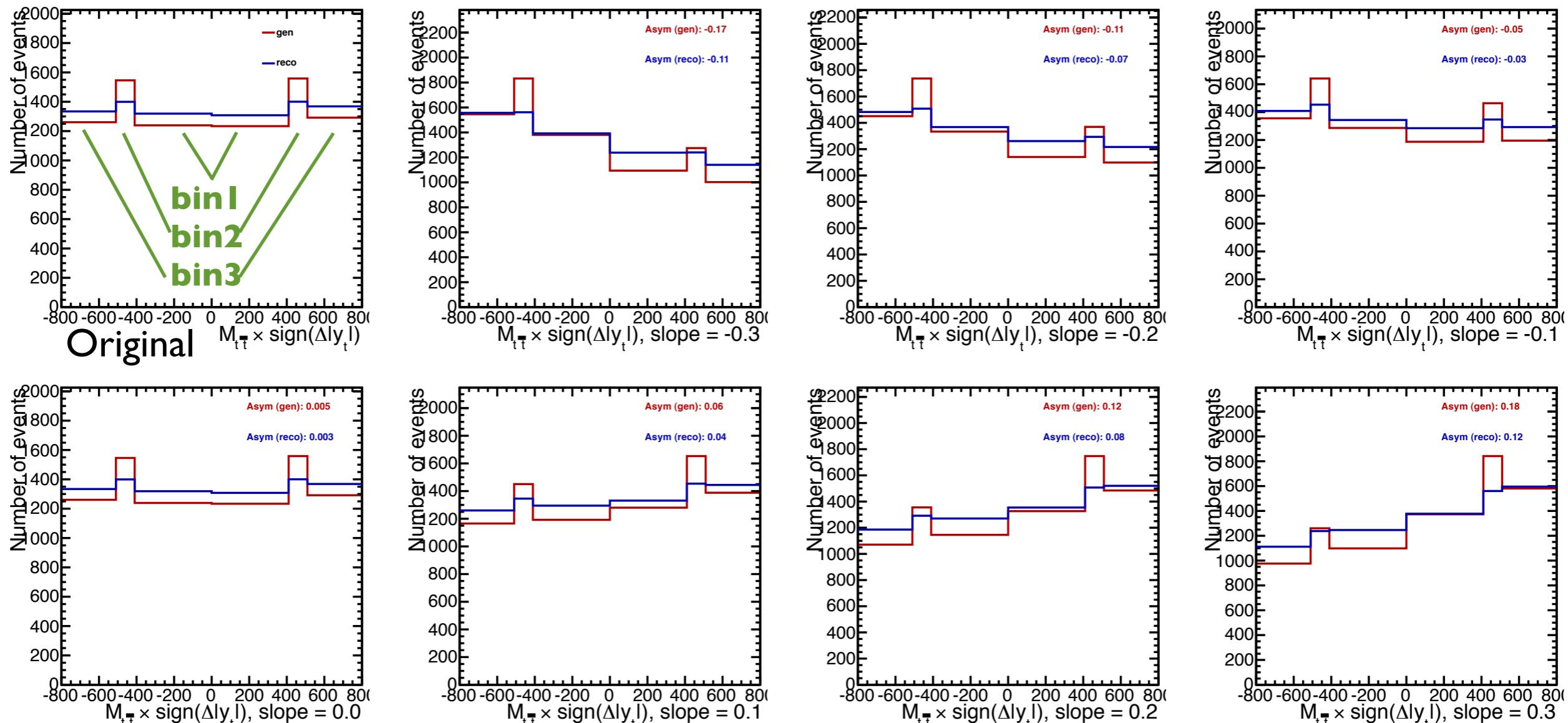
- In TOP-12-010 we are measuring the top and lepton charge asymmetries,  $A_C$  and  $A_{lepC}$ , both inclusively and their differential dependence on  $M_{ttbar}$ ,  $P_{T,ttbar}$ , and  $y_{ttbar}$
- We've found a problem with the linearity of our unfolding procedure for the  $A_C$  differential measurements
  - it is related to our method of converting the “2D” unfolding into 1D by multiplying  $M_{ttbar}$ ,  $P_{T,ttbar}$ , or  $|y_{ttbar}|$  by the sign of the asymmetry variable (as done by CDF), meaning we effectively have only 2 bins in the asymmetry variable
- These slides focus on the  $M_{ttbar}$  dependence, but the results and conclusions are the same for the  $P_{T,ttbar}$  and  $y_{ttbar}$  dependence

# Linearity test method

- We add asymmetry to the default MC distribution by applying weight=1+slope\* $\Delta|y_{\text{gen}}|$  per event (see plots below), then compare unfolded asymmetry to true asymmetry, inclusively and in the 3 bins we use for the differential measurement

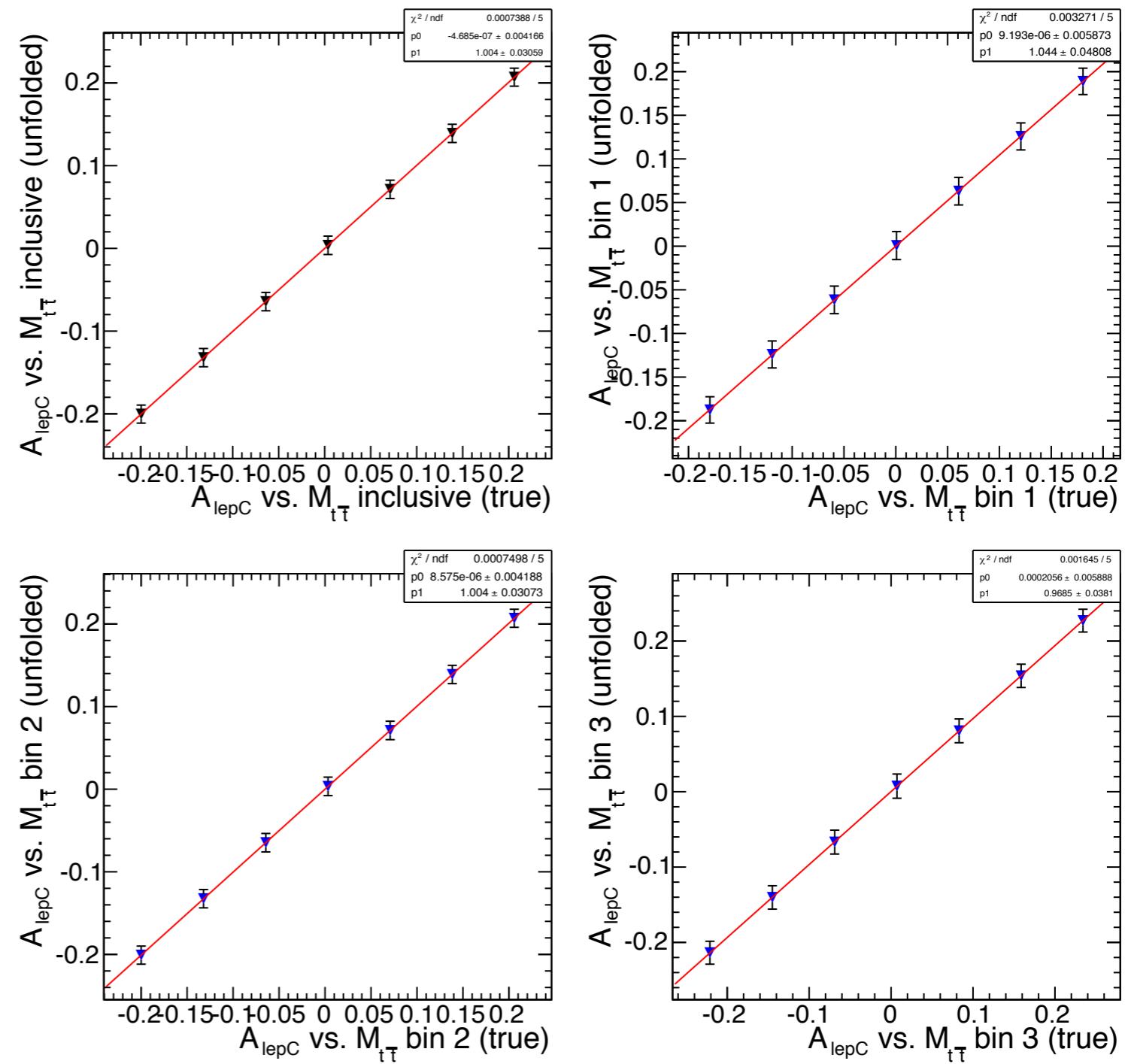


## Original and weighted distributions for $M_{t\bar{t}} * \text{sign}(\Delta|y|)$ , before unfolding:



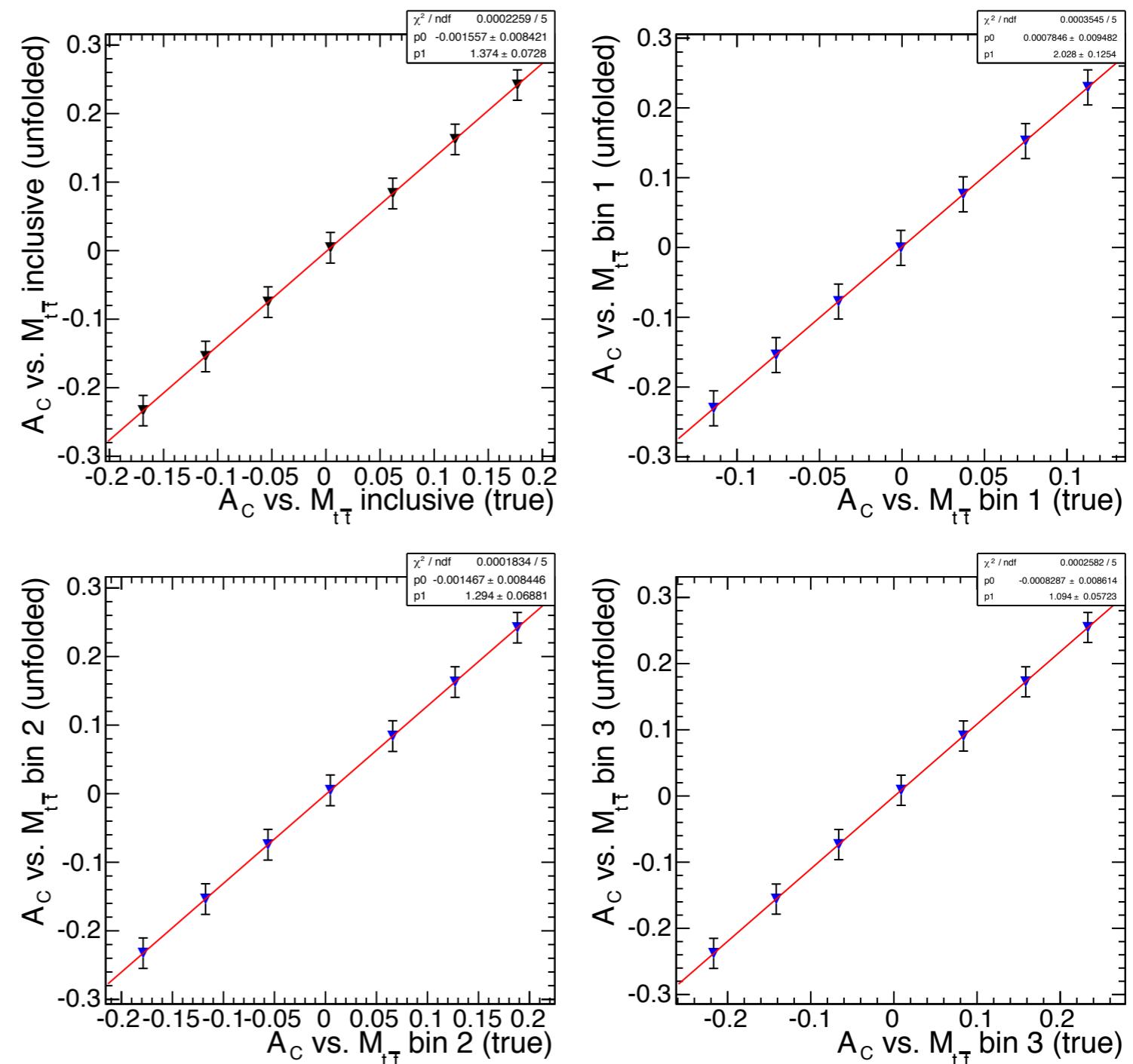
# Results for $A_{\text{lepC}}$

- Linearity looks pretty good for all bins ( $p \mid \sim \mid$ )



# Results for $A_C$

- Things don't look so good here
  - Slope of linearity plot (p1) significantly away from one for bin1 (2.03) and bin2 (1.29), as well as inclusively (1.37)
- We think this is an inherent bias of using 2 bins in  $\Delta|y|$  to unfold distributions with a continuous slope in  $\Delta|y_{gen}|$  (as expected for a real physics effect, and which we generate using weight = 1+slope\* $\Delta|y_{gen}|$ , see slide 3)

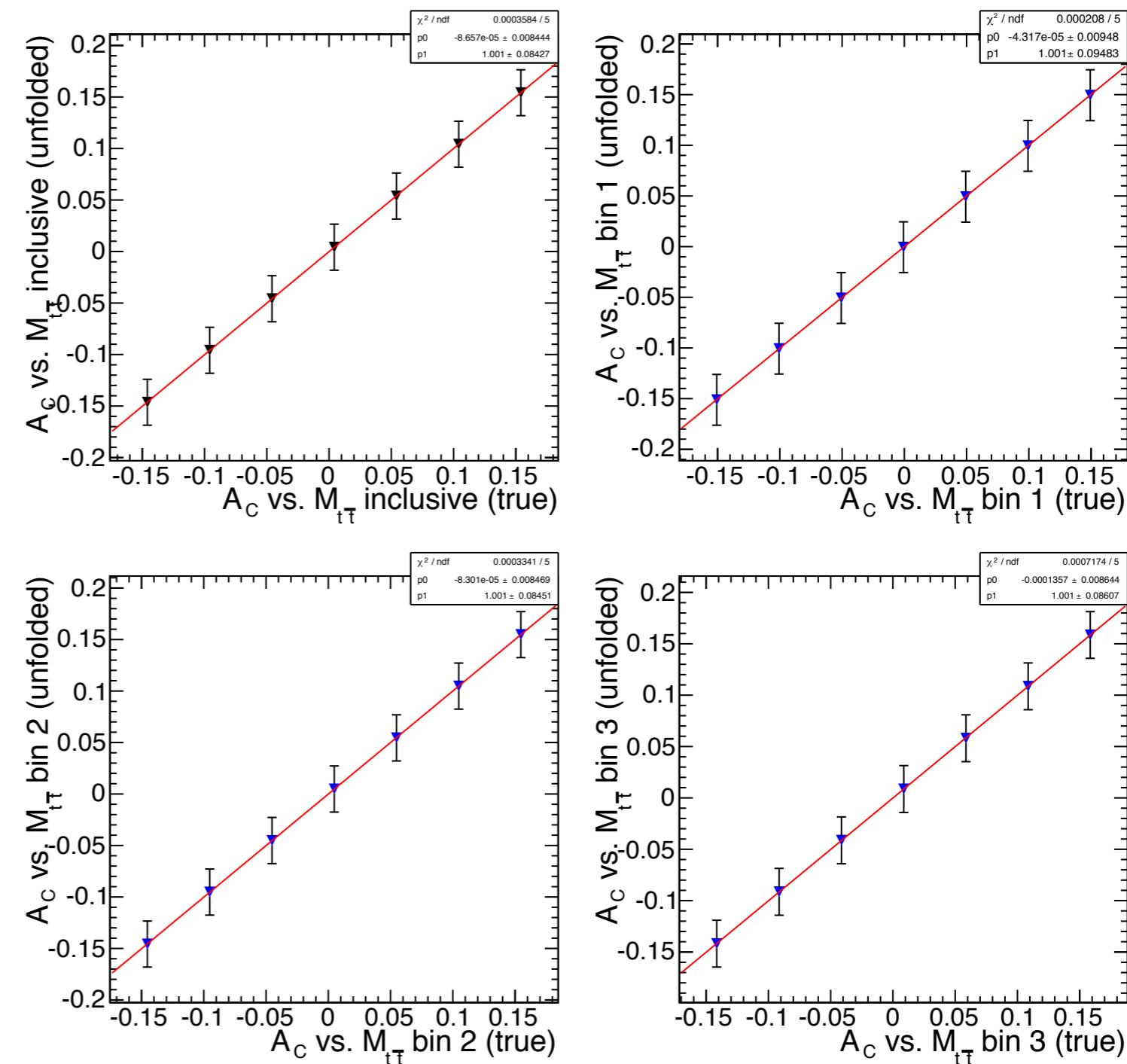


# Bias from insufficient $\Delta|y|$ bins

- We add a slope by applying weight =  $1 + \text{slope} * \Delta|y_{\text{gen}}|$ 
  - => events with large  $|\Delta|y_{\text{gen}}|$  get the largest change in weight
    - but events with large  $|\Delta|y_{\text{gen}}|$  are also the events least prone to changing sign of  $\Delta|y_{\text{gen}}|$  through migration
    - however, the response matrix treats them as though they had the average chance of migration => biased unfolding
  - With a slope matching the granularity of the response matrix, i.e. weight  $1 + 0.5 * \text{slope} * \text{sign}(\Delta|y_{\text{gen}}|)$ , weighting affects all events equally
    - => should be no bias
    - this is confirmed on the next slide

# Results for $A_C$ (slope only in $\text{sign}(\Delta|y_t|)$ )

- Slope now very close to  $p1=1$  for all bins
- this result holds for both asymmetry variables, and unfolding vs  $M_{t\bar{t}}$ ,  $pT_{t\bar{t}}$  and  $y_{t\bar{t}}$
- but of course the results on slides 4 and 5 are a more accurate estimate of the linearity of the method for a general physics effect
- The only reason we can get away with using 2  $\Delta|y|$  bins for  $A_{lepC}$  (slide 4) is that the response matrix basically does nothing here (excellent experimental resolution)



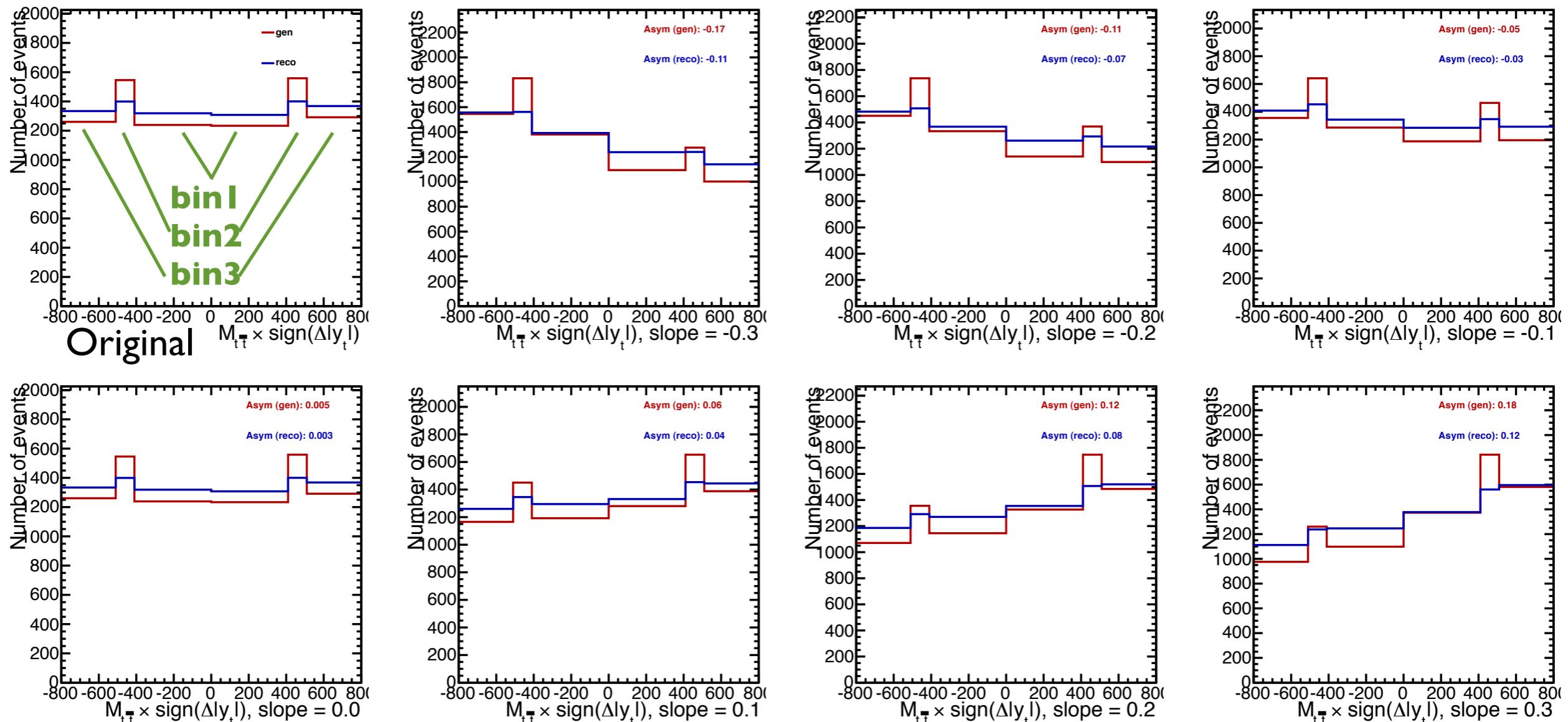
# Conclusions

- Linearity of the method is pretty good for the differential dependence of  $A_{\text{lepc}}$ , but poor for  $A_c$ 
  - bias comes from using a response matrix with insufficient resolution (only 2 bins in  $\Delta|y|$ )
- We think showing 2D results for just 1 of the 2 variables might cause problems getting the paper published
- We propose to include only the 1D results in the paper, i.e. the values and distributions for  $A_c$  and  $A_{\text{lepc}}$

# Backup

# Weighted distributions with continuous slope in $\Delta|y_t|$

- plots show original results with weight =  $|+slope * \Delta|y_{gen}|$
- Inclusive **reco-level asym** is  $\sim 2/3 * \text{gen-level asym}$ , prior to unfolding (see numbers on plots)



# Weighted distributions with slope in $\text{sign}(\Delta|y_t|)$

- weight  $1 + 0.5 * \text{slope} * \text{sign}(\Delta|y_{\text{gen}}|)$
- **Reco-level asym** is now only  $\sim 1/2 * \text{gen-level asym}$ , prior to unfolding
  - => reco distribution responds very differently to the different types of slope, but the effect of applying our “2-bin” response matrix will be the same in both cases
  - => expect pI (inclusive) to be  $(1/2)/(2/3) * 1.37 \approx 1.0$  when we add a slope in  $\text{sign}(\Delta|y_t|)$  instead of  $\Delta|y_t|$ , as is observed on slides 5 and 7

